



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113874

TO: Manjunath N Rao
Location: REM/3B81/3C70
Art Unit: 1652
February 18, 2004

Case Serial Number: 10059447

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:29:41 ; Search time 11 Seconds

(without alignments)
123.979 Million cell updates/sec

Title: US-10-059-447A-1

Perfect score: 153

Sequence: 1 MKVLGNYIQRMFYDGKXYTKQFNKPIKX 29

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	32.7	651	1	SYM METH
2	49.5	32.4	1582	1	YU30_RALSO
3	49	32.0	521	1	CP51_CANPA
4	48	31.4	579	1	ILVB_SPLPL
5	47.5	31.0	645	1	LIP1_PHOUL
6	47	30.7	80	1	R27A_MOUSE
7	46.5	30.4	1516	1	Y819_PSEAB
8	46	30.1	271	1	NOR1_ASPPA
9	45	30.1	2514	1	POINTONVPA
10	45.5	29.7	530	1	FMO3_RABIT
11	45.5	29.7	623	1	HTPG_CLOPE
12	45.5	29.7	633	1	YR45_CAEEL
13	45	29.4	255	1	THID_AQUAE
14	45	29.4	256	1	TPIS_VIBPA
15	45	29.4	302	1	SC14_CANGA
16	45	29.4	350	1	DHVS_SCHPO
17	45	29.4	512	1	EMRB_ECOLI
18	45	29.4	763	1	PEPX_LACLA
19	45	29.4	1319	1	SOSI_MOUSE
20	45	29.4	1333	1	SOSI_HUMAN
21	44.5	29.1	317	1	PLC_LISMO
22	44.5	29.1	864	1	IF2_BUCAL
23	44	28.8	95	1	Y060_NPYLD
24	44	28.8	98	1	SY19_HUMAN
25	44	28.8	185	1	KITH_CPEPV
26	44	28.8	186	1	KITH_CPEPV
27	44	28.8	343	1	GUNC_CLOSF
28	44	28.8	343	1	GUNC_CLOSF
29	44	28.8	601	1	ISPG_CHLMU
30	44	28.8	772	1	LPIG_DROME
31	44	28.8	802	1	CSD1_ECOLI
32	44	28.8	819	1	CSD1_ECOLI
33	44	28.8	1348	1	VGR2_COTUA

ALIGNMENTS

RESULT 1	SYM METH	STANDARD	PRT	651 AA
AC	026587			
DT	15-0UL-1998 (Rel. 36, Created)			
DT	15-0UL-1998 (Rel. 36, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MecRS)			
OS	MENG OR MTH587			
OC	Methanobacterium thermoautotrophicum			
OC	Archaea; Buryarchaeota; Methanobacteria; Methanobacteriales;			
OC	Methanobacteriaceae; Methanothermobacter			
OX	NCBI_TaxID=187420;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Delta H;			
RX	MEDLINE=98037514; PubMed=9371463;			
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,			
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,			
RA	Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,			
RA	Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,			
RA	Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,			
RA	McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,			
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;			
RT	"Complete genome sequence of Methanobacterium thermoautotrophicum			
RT	deltah: functional analysis and comparative genomics."/;			
RL	J. Bacteriol. 179:7135-7155(1997).			
CC	-1- FUNCTION: Is required not only for elongation of protein synthesis			
CC	but also for the initiation of all mRNA translation through			
CC	initiator tRNA(Met) aminoacylation (By similarity).			
CC	-1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +			
CC	diphosphate + L-methionyl-tRNA(Met).			
CC	-1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).			
CC	-1- SUBUNIT: Homodimer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.			
CC	MecG subfamily 1.			
CC	-1- SIMILARITY: Contains 1 tRNA-binding domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL; AE000841; AAB85093.1; -			
DR	PIR; G69177; G69177.			
DR	HSSP; P00959; IMEA.			
DR	HAMAP; MF_00098; fused; 1.			
DR	InterPro; IPR004495; MetG_Cterm.			
DR	InterPro; IPR002300; tRNA-synt_1a.			
DR	InterPro; IPR001412; tRNA-synt_1.			
DR	InterPro; IPR002304; tRNA-synt_mec.			

DR InterPro; IPR002547; tRNA bind.
 DR Pfam; PF00133; tRNA-lyt1; 1.
 DR Pfam; PF01588; tRNA bind. 1.
 DR PIRSF; PIRSF001528; tRNA dimerising; 1.
 DR PRINTS; PR01041; TRNASYNTHET.
 DR TIGRfam; TIGR00398; metG_1.
 DR TIGRfam; TIGR00399; metG_C-term; 1.
 DR PROSITE; PS00178; AA tRNA_LIGASE_1; FALSE_NEG.
 DR PROSITE; PS50886; TRD; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM RNA-binding; tRNA-binding; Metal-binding; zinc; Complete proteome.
 FT SITE 11 21
 FT SITE 325 329 "HIGH" REGION.
 FT SITE 554 651 "KSKS" REGION.
 FT DOMAIN 143 143 tRNA-BINDING.
 FT METAL 143 143 ZINC (BY SIMILARITY).
 FT METAL 146 146 ZINC (BY SIMILARITY).
 FT METAL 156 156 ZINC (BY SIMILARITY).
 FT METAL 159 159 ZINC (BY SIMILARITY).
 SQ SEQUENCE 651 AA; 74631 MW; ABA747D9E7DBE2C CRC64;

Query Match 32.7%; Score 50; DB 1; Length 651;
 Best Local Similarity 47.1%; Pred. No. 7.8;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 VLGNVIGRNHYDGKFF 19
 Db 380 VLGNVIGRNHYDGKFF 386

RESULT 2
 ID YJ30 RALSO STANDARD; PRT; 1582 AA.
 AC OXKX02;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RSC3030 precursor.
 GN RSC3030 OR R504727.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; Pubmed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangent S.,
 RA Atlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Gaspin C., Lavie W., Molien A., Robert C., Saurin W., Schlex T.,
 RA Signer P., Thepaut P., Whalen M., Winkler P., Levy N.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL646073; CADI6739.1; -
 DR KJ000000; HYPOHETICAL PROTEIN RSC3030.
 FT SIGNAL 15
 FT CHAIN 16 1582
 SQ SEQUENCE 1582 AA; 170090 MW; 8683D148F5A3C2A CRC64;

Query Match 32.4%; Score 49.5; DB 1; Length 1582;
 Best Local Similarity 33.3%; Pred. No. 23;
 Matches 9; Conservative 8; Mismatches 7; Indels 3; Gaps 1;

Qy 5 GNVIGRNHYDGKFFTKCF---NKPI 28
 Db 316 GNVIGRNHYDGKFFTKCF---NKPI 342

RESULT 3
 ID CP51 CANMA STANDARD; PRT; 521 AA.
 AC Q12586;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 52A9 (BC 1.14.14.-) (CYPL1A9) (Alkane-inducible
 DE P450-ALKS-A).
 DE CYP52A9.
 OS Candida maltosa (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5479;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95169275; Pubmed=7865134;
 RA Ohkuma M., Maeda S., Tanimoto T., Fujii M., Ohta A., Takagi M.;
 RA "CYP52 (cytochrome P450alk) multigene family in Candida maltosa:
 RT identification and characterization of eight members";
 RL DNA Cell Biol. 14:163-173(1995).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=96311366; Pubmed=8713123;
 RA Zimmer T., Ohkuma M., Ohta A., Takagi M., Schunck W.H.;
 RA "The CYP52 multigene family of Candida maltosa encodes functionally
 RT diverse n-alkane-inducible cytochromes P450";
 RL Biochem. Biophys. Res. Commun. 224:784-789(1996).
 CC -1- FUNCTION: TOGETHER WITH AN NADPH CYTOCHROME P450 THE ENZYME SYSTEM
 CC CATALYZES THE TERMINAL HYDROXYLATION AS THE FIRST STEP IN THE
 CC ASSIMILATION OF ALKANES AND FATTY ACIDS.
 CC -1- CATALYTIC ACTIVITY: Hydroxylation of N-alkanes at the terminal
 CC position.
 CC -1- INDUCTION: By N-alkanes.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D12717; BAA02211.1; -
 DR PIR; J50723; J50723.
 DR HSSP; P14779; IUPZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.
 FT METAL 468
 FT METAL 468 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 521 AA; 59872 MW; 07D8827DC92042DA CRC64;

Query Match 32.0%; Score 49; DB 1; Length 521;
 Best Local Similarity 40.0%; Pred. No. 8.8;
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 7 YIGRNHYDGKFFTKCF 26
 Db 217 YIGRNHYDGKFFTKCF 236

RESULT 4
 ID ILVB_SPIPL STANDARD; PRT; 579 AA.
 ID ILVB_SPIPL

```

AC P27868:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Acetylacetyl synthase (EC 2.2.1.6) (Acetylhydroxy-acid synthase) (ALS)
DE (fragment)
GN ILVY.
OS Spirulina platensis.
OC Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
OX NCBI_TaxID=1156;
RN [1]
RX MEDLINE=92381487; PubMed=1512571;
RA Milano A., de Rossi E., Zanaria E., Barbierato L., Ciferri O.,
RA Ricciardi G.;
RT "Molecular characterization of the genes encoding acetylhydroxy acid
RT synthase in the cyanobacterium Spirulina platensis."
RL J. Gen. Microbiol. 138:1399-1408(1992).
CC -1- CATALYTIC ACTIVITY: 2 pyruvate = 2-acetylacetyl + CO(2).
CC -1- COFACTOR: Binds 1 thiamine pyrophosphate and 1 magnesium ion per
CC subunit (By similarity).
CC -1- PATHWAY: Valine and isoleucine biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE TPP ENZYMES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75907; AAA26595.1; -
DR PIR; B44857; B44857.
DR HSSP; P06169; 1PVD.
DR InterPro: IPR004407; AcCoA19.
DR InterPro: IPR000399; Pyruvate_decarb.
DR Pfam; PF00205; TPP_enzymes; 1.
DR Pfam; PF02775; TPP_enzymes; 1.
DR Pfam; PF02776; TPP_enzymes; 1.
DR TIGRfam; TIGR00118; acCoA19; 1.
DR PROSITE; PS00187; TPP_ENZYMES; PARTIAL.
KW Transferase; Branched-chain amino acid biosynthesis; Flavoprotein;
KW Thiamine pyrophosphate.
FT ACT_SITE 61
FT NON_TER 579
SQ SEQUENCE 579 AA; 63440 MW; BC94FEA728A7889A CRC64;

Query Match 31.4%; Score 48; DB 1; Length 579;
Best Local Similarity 31.4%; Pred. No. 14;
Matches 11; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY 2 KVLGNYIQNPFHYDGKXFFYTK-----QFNKPI 28
DB 339 QVLGDMVQRTYHWERTLSRKPRNGTDLNQLREPI 373

RESULT 5
LIP1_PRODU STANDARD; PRT; 645 AA.
ID LIP1_PRODU STANDARD; PRT; 645 AA.
AC P40601;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lipase 1 precursor (EC 3.1.1.3) (Triacylglycerol lipase).
GN LIP-1.
OS Photorhabdus luminescens (Xenorhabdus luminescens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=29488;
RN [1]
RX MEDLINE=93194791; PubMed=8449874;
RA Wang H., Dows B.C.A.;
RT "Phase variation in Xenorhabdus luminescens: cloning and sequencing
RT of the lipase gene and analysis of its expression in primary and
RT secondary phases of the bacterium."
RL J. Bacteriol. 175:1665-1673(1993).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66379; CAA47020.1; -
DR PIR; A47081; A47081.
DR InterPro: IPR006315; Autotransport.
DR InterPro: IPR005546; Autotransporter.
DR InterPro: IPR001087; Lipase_GDSL.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF00657; Lipase_GDSL; 1.
DR TIGRfam; TIGR01414; Autotrans bar1; 1.
DR PROSITE; PS01098; LIPASE_GDSL_SER; 1.
KW Hydrolase; Lipid degradation; Signal.
FT SIGNAL 1
FT CHAIN 24
FT ACT_SITE 25 645 LIPASE 1.
FT ACT_SITE 34 34 BY SIMILARITY.
FT ACT_SITE 330 330 POTENTIAL.
SQ SEQUENCE 645 AA; 70716 MW; 1271327BC56932F CRC64;

Query Match 31.0%; Score 47.5; DB 1; Length 645;
Best Local Similarity 38.5%; Pred. No. 18;
Matches 10; Conservative 3; Mismatches 4; Indels 9; Gaps 1;

QY 3 VLGNVIR-----NFHYDGKX 19
DB 415 LIGCMISRYQDNSSPANNFYHDGNGY 440

RESULT 6
R27A_MOUSE STANDARD; PRT; 80 AA.
ID R27A_MOUSE STANDARD; PRT; 80 AA.
AC P49664; Q9D2W3;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 40S ribosomal protein 527a.
GN RPS27A.
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090; 10116;
RN [1]
RX MEDLINE=96011832; PubMed=7488009;
RA Chan Y.-L., Suzuki K., Wool I.G.;
RT "The carboxyl extensions of two rat ubiquitin fusion proteins are
RT ribosomal proteins 527a and 140."
RL Biochem. Biophys. Res. Commun. 215:682-690(1995).
RN [2]
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,

```

```

RX MEDLINE=93194791; PubMed=8449874;
RA Wang H., Dows B.C.A.;
RT "Phase variation in Xenorhabdus luminescens: cloning and sequencing
RT of the lipase gene and analysis of its expression in primary and
RT secondary phases of the bacterium."
RL J. Bacteriol. 175:1665-1673(1993).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66379; CAA47020.1; -
DR PIR; A47081; A47081.
DR InterPro: IPR006315; Autotransport.
DR InterPro: IPR005546; Autotransporter.
DR InterPro: IPR001087; Lipase_GDSL.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF00657; Lipase_GDSL; 1.
DR TIGRfam; TIGR01414; Autotrans bar1; 1.
DR PROSITE; PS01098; LIPASE_GDSL_SER; 1.
KW Hydrolase; Lipid degradation; Signal.
FT SIGNAL 1
FT CHAIN 24
FT ACT_SITE 25 645 LIPASE 1.
FT ACT_SITE 34 34 BY SIMILARITY.
FT ACT_SITE 330 330 POTENTIAL.
SQ SEQUENCE 645 AA; 70716 MW; 1271327BC56932F CRC64;

Query Match 31.0%; Score 47.5; DB 1; Length 645;
Best Local Similarity 38.5%; Pred. No. 18;
Matches 10; Conservative 3; Mismatches 4; Indels 9; Gaps 1;

QY 3 VLGNVIR-----NFHYDGKX 19
DB 415 LIGCMISRYQDNSSPANNFYHDGNGY 440

RESULT 6
R27A_MOUSE STANDARD; PRT; 80 AA.
ID R27A_MOUSE STANDARD; PRT; 80 AA.
AC P49664; Q9D2W3;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 40S ribosomal protein 527a.
GN RPS27A.
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090; 10116;
RN [1]
RX MEDLINE=96011832; PubMed=7488009;
RA Chan Y.-L., Suzuki K., Wool I.G.;
RT "The carboxyl extensions of two rat ubiquitin fusion proteins are
RT ribosomal proteins 527a and 140."
RL Biochem. Biophys. Res. Commun. 215:682-690(1995).
RN [2]
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,

```

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gasterland T., Gissi C., King B., Kocich H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pessio G., Quackenbush J.,
 RA Sakai K., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Boffelli D., Furuno M., Kono H., Balderelli R., Barsh G.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).

[3] SEQUENCE FROM N.A.

RP SPECIES=mouse;
 RC MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rask S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millhys S.J.,
 RA Bask S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heaton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marz M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

-1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
 C-TERMINAL EXTENSION PROTEIN (CEP) OF UBQUITIN.

-1- SIMILARITY: BELONGS TO THE S27AE FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EXBL: X81839; CAA57432.1; ALT INIT.
 DR EMBL: AK018706; BAB31357.1; ALT INIT.
 DR EMBL: BC002108; AA02108.1; ALT INIT.
 DR HSSP: P02246; IUBI.
 DR WGD: MGI:1925544; Rpe27a.
 DR InterPro: IPR002906; Ribosomal S27.
 DR Pfam: PF01599; Ribosomal S27; 1.
 KM Ribosomal protein, zinc-finger; Metal-binding.
 FT DOMAIN 1 23 LYS-RICH (HIGHLY BASIC).
 FT ZN FING 45 68 C4-TYPE
 SQ SEQUENCE 80 AA; 9404 MW; F255F54B6F9DC1F CRC64;

Query Match 30.7%; Score 47; DB 1; Length 80;
 Best Local Similarity 45.8%; Pred. No. 2.7;
 Matches 11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

OY 4 LGNYIQNFHYDGKFFYTKQFNP 27
 DB 56 GMSHPDR--HYCGKCLITCFNKP 77

RESULT 7

Y819_PSEAE

ID Y819_PSEAE

STANDARD;

PRT; 1516 AA.

AC OSHV2;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein PA4489 precursor.

GN PA4489.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Brody L.L., Golety L., Tolentino E., Westbrock-Nadman S., Yuan Y.,
 RA Smith K.A., Spencer D.H., Wong K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).

-1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE004862; AA007877.1; -

DR PIR: F83085; F83085.

KW Hypothetical protein; Signal; Complete proteome.

FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 1516 HYPOTHEITICAL PROTEIN PA4489.

FT SEQUENCE 1516 AA; 167428 MW; FDCED1561D62C18C CRC64;

Query Match 30.4%; Score 46.5; DB 1; Length 1516;
 Best Local Similarity 29.6%; Pred. No. 60;
 Matches 8; Conservative 9; Mismatches 7; Indels 3; Gaps 1;

OY 5 GNYIQNFHYDGKFFYTKQF---NKPI 28
 DB 315 GVFSNPFYDSIYDKRIYATDRPL 341

RESULT 8

NOR1_ASPPA

ID NOR1_ASPPA

STANDARD;

PRT; 271 AA.

AC Q00278;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Aflatoxin biosynthesis ketoreductase NOR-1 (EC 1.1.1.-).

GN NOR-1 OR NAR-1.

OS Aspergillus parasiticus.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxID=5067;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 163 / NRRL 5862 / SU-1;

RX MEDLINE=95085270; PubMed=7993094;

RA Trall F., Chang P.-K., Cary J., Linz J.E.;
 RT "Structural and functional analysis of the nor-1 gene involved in the
 RT biosynthesis of aflatoxins by Aspergillus parasiticus.";

```

RL Appl. Environ. Microbiol. 60:4078-4085(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE DEHYDRATION OF NORSOLONIC ACID
CC TO FORM AVERANTIN.
CC -1- PATHWAY: Aflatoxin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY. STRONG, TO E.NIDULANS SPEC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L27801; AA58798.1; -
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
KW Oxidoreductase
FT NP_BIND 26 52 NAD OR NADP (BY SIMILARITY).
FT ACT_SITE 185 185 BY SIMILARITY.
SQ SEQUENCE 271 AA; 29569 MW; 34059A857672059A CRC64;
Query Match 30.1%; Score 46; DB 1; Length 271;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 KVLGNYIQNNPHYDGK 17
Db 189 KLANVYVRKHFENK 204

RESULT 9
POINT ONVNG STANDARD; PRT; 2514 AA.
AC P1386;
DT 01-JUN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nonstructural polypeptide [Contains: Nonstructural protein NSP1;
DE Nonstructural protein NSP2; Nonstructural protein NSP3; Nonstructural
DE protein NSP4].
OS O'nyong-nyong virus (strain Gulu) (ONNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_Taxid=11028;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90177206; PubMed=2155505;
RA Levinson R.S., Strauss J.H., Strauss E.G.,
RT "Complete sequence of the genomic RNA of O'nyong-nyong virus and its
RT use in the construction of alphavirus phylogenetic trees.";
RL Virology 175:110-123(1990).
RN [2]
RP SEQUENCE OF 1334-2514 FROM N.A.
RX MEDLINE=88206074; PubMed=2834873;
RA Strauss E.G., Levinson R., Rice C.M., Dalrymple J., Strauss J.H.,
RT "Nonstructural proteins nspp3 and nspp4 of Ross River and O'nyong-nyong
RT viruses: sequence and comparison with those of other alphaviruses.";
RL Virology 164:265-274(1988).
RN [3]
RP FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20303; AAA46784.1; -

```

```

DR PIR; A34680; MNWVW2.
DR MEROPS; C09.001; -.
DR InterPro; IPR002589; Aldp.
DR InterPro; IPR002620; peptidase C9.
DR InterPro; IPR001788; RNA_dep_RNApol2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF01661; Aldp; 1.
DR Pfam; PF01707; peptidase C9; 1.
DR Pfam; PF00978; RNA_dep_RNApol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR SMART; SMO0505; Aldp; 1.
KW Polypeptide; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 535 NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 536 1333 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1334 1903 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1904 2514 NONSTRUCTURAL PROTEIN NSP4.
SQ SEQUENCE 2514 AA; 280114 MW; 3708374690530E21 CRC64;
Query Match 30.1%; Score 46; DB 1; Length 2514;
Best Local Similarity 38.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 8 IQBNHYDGKKEFYTKQFNKPI 28
Db 855 IVSSLHYSKRTNRYNPI 875

RESULT 10
FM03 RABIT STANDARD; PRT; 530 AA.
AC P32417;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylamine monooxygenase [N-oxide forming] 3 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 3) (FMO 3) (Dimethylamine
DE oxidase 3) (FMO ID1) (FMO form 2) (FMO II).
GN FM03.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=New Zealand white; TISSUE=Liver;
RX MEDLINE=94245761; PubMed=8188717;
RA Burnett V.L., Lawton M.P., Philpott R.M.,
RT "Cloning and sequencing of flavin-containing monooxygenases FMO3 and
RT FMO4 from rabbit and characterization of FMO3.";
RL J. Biol. Chem. 269:14314-14322(1994).
RN [2]
RP SEQUENCE.
RX MEDLINE=91378577; PubMed=1898080;
RA Ozols J.,
RT "Multiple forms of liver microsomal flavin-containing monooxygenases:
RT complete covalent structure of form 2.";
RL Arch. Biochem. Biophys. 290:103-115(1991).
RN [3]
RP SEQUENCE OF 1-32.
RC TISSUE=Liver;
RX MEDLINE=89374273; PubMed=2505769;
RA Ozols J.;
RT "Liver microsomes contain two distinct NADPH-Monooxygenases with NH2-
RT terminal segments homologous to the flavin containing NADPH-
RT monooxygenase of Pseudomonas fluorescens.";
RL Biochem. Biophys. Res. Commun. 163:49-55(1989).
CC -1- FUNCTION: Involved in the oxidative metabolism of a variety of
CC xenobiotics such as drugs and pesticides. It N-oxygenates primary
CC aliphatic alkyamines as well as secondary and tertiary amines.
CC -1- CATALYTIC ACTIVITY: N,N-dimethylamine + NADPH + O(2) = N,N-
CC dimethylamine N-oxide + NADP(+) + H(2)O.

```

```

CC -1- COFACTOR: FAD.
CC -1- SUBCELLULAR LOCATION: Microsomal.
CC -1- TISSUE SPECIFICITY: Liver.
CC -1- SIMILARITY: BELONGS TO THE FMO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: L10391; AAA21178.1; -.
CC PIR: B54096; B54096.
CC PIR: S18380; S18380.
CC InterPro: IPR001327; FAD_Pyr_redox.
CC InterPro: IPR000960; Flav_cont_mnoxgn.
CC InterPro: IPR001100; Pyr_redox.
CC Pfam: PF00743; FMO-like; 1.
CC PRINTS: PR00368; FADPFR.
CC PRINTS: PR00370; FMOXYGENASE.
CC PRINTS: PR00411; ENDRDASEL.
CC Oxidoreductase; Monooxygenase; NADP; Flavoprotein; FAD; Microsome;
CC Transmembrane; Multigene family; Polymorphism.
CC INIT_MER 0 0
CC NP_BIND 8 13 FAD (ADP PART) (POTENTIAL).
CC NP_BIND 190 195 NADP (POTENTIAL).
CC VARIANT 278 278 R -> M.
CC VARIANT 404 404 M -> V.
CC CONFLICT 75 75 D -> P (IN REF. 2).
CC CONFLICT 80 80 F -> N (IN REF. 2).
CC CONFLICT 127 129 STE -> ATC (IN REF. 2).
CC CONFLICT 172 172 HS -> RQ (IN REF. 2).
CC CONFLICT 196 196 C -> E (IN REF. 2).
CC CONFLICT 305 305 F -> EXEF (IN REF. 2).
CC CONFLICT 418 418 W -> T (IN REF. 2).
CC CONFLICT 422 422 S -> W (IN REF. 2).
CC CONFLICT 513 515 WLK -> ELM (IN REF. 2).
CC SEQUENCE 530 AA; 59684 MW; A450407620775CC8 CRC64;

Query Match 29.7%; Score 45.5; DB 1; Length 530;
Best local Similarity 21.9%; Pred. No. 30;
Matches 7; Conservative 10; Mismatches 6; Indels 9; Gaps 1;

QY 5 GNYIQNKF-----HYDGKXFTYKQFNKP 27
DB 147 GHIVYENLPKDSPPGUKHFKGKSPHSRYEKPE 178

RESULT 11
HTPG_CLOPE STANDARD; PRT; 623 AA.
AC O8XNC2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature
DE protein G)
GN HTPG OR CPE0415.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1502;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A.
RX MEDLINE=21664373; Pubmed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Onshima K., Yamashita A.,
RA Shiba T., Ogassawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

```

```

CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AP003186; BAB80122.1; -.
CC HAMAP: MF_00505; -1.
CC InterPro: IPR003594; ATPbind_ATPase.
CC InterPro: IPR001404; Hsp90.
CC Pfam: PF02518; HATPase_c/1.
CC Pfam: PF00183; HSP90; 1.
CC PRINTS: PR00775; HEATSHOCK90.
CC SMART: SM00387; HATPase_c/1.
CC PROSITE: PS00298; HSP90; FALSE NEG.
CC Chaperone; ATP-binding; Heat shock; Complete proteome.
CC DOMAIN 1 341 A; SUBSTRATE-BINDING (BY SIMILARITY).
CC DOMAIN 342 549 B (BY SIMILARITY).
CC DOMAIN 550 623 C.
CC SEQUENCE 623 AA; 72451 MW; 99DD4F2E8BEC652F CRC64;

Query Match 29.7%; Score 45.5; DB 1; Length 623;
Best local Similarity 45.0%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY 7 YIQNFYDGKXFTYKQFNKP 26
DB 284 YIPENIYD---YITKYEYK 300

RESULT 12
YR45_CABEL STANDARD; PRT; 633 AA.
AC Q09562;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 72.9 kDa protein P47D12.5 in chromosome III.
GN P47D12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Tatch A.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: WEAK, TO C.ELEGANS ZYG-11.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U22831; AAK20066.1; -.
CC PIR: P88478; P88478.
CC Wormpep; P47D12.5; CE01949.
CC Hypothetical protein.
CC SEQUENCE 633 AA; 72904 MW; A205565802E608FD CRC64;

Query Match 29.7%; Score 45.5; DB 1; Length 633;

```


Best Local Similarity 34.5%; Pred. No. 35;
Matches 10; Conservative 6; Mismatches 10; Indels 3; Gaps 1;

QY 3 VLGNVIOBNHY--DCKXYTKQFNKPI 28
DB 41 INNTERRNHLYIDIKGSHAFSKMTNTI 69

RESULT 13

THID_AQUAE STANDARD; PRT; 255 AA.

AC 067772;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphomethylpyrimidine kinase (EC 2.7.4.7) (HMP-phosphate kinase)
GN THID OR AQ_1960.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=V95;
RX MEDLINE=9816666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus";
RL Nature 392:353-358(1998).

CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF HMP-P TO HMP-PP (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 4-amino-2-methyl-5-
CC phosphomethylpyrimidine = ADP + 4-amino-2-methyl-5-
CC diposphomethylpyrimidine.
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE THID FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE000765; AAC07733.1; -.
DR PIR; H70467; H70467.
DR InterPro; IPR004339; HMP-P_kinase.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00294; PfkB; 1.
DR TIGRfam; TIGR00097; HMP-P_kinase; 1.
KW Thiamine biosynthesis; transferase; kinase; complete proteome.
SQ SQUONCE 255 AA; 28065 MW; 9544CASF662F708F CRC64;

Query Match 29.4%; Score 45; DB 1; Length 255;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 6; Mismatches 7; Indels 6; Gaps 2;

QY 3 VLGNVIOBNHY--YDCKXYTKQFNKPI 27
DB 171 IKGHLKGNVAIDILYKSFY--EFKAP 197

RESULT 14

TPIS_VIBRA STANDARD; PRT; 256 AA.

AC 067751;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).

GN TPIS OR VP0239.

OC Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kibota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae";
RL Lancet 361:743-749(2003).

CC -1- CATALYTIC ACTIVITY: D-glyceridehyde 3-phosphate = glyceroe
CC phosphate.
CC -1- PATHWAY: Plays an important role in several metabolic pathways.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AP005073; BAC58502.1; -.
DR HAMAP; MF_00147; -1.
DR PROSITE; PS00171; TIM; 1.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt; Complete proteome.
FT ACT SITE 97 97 BY SIMILARITY.
FT ACT SITE 169 169 BY SIMILARITY.
SQ SEQUENCE 256 AA; 26932 MW; 50A14E37B34CCAF7 CRC64;

Query Match 29.4%; Score 45; DB 1; Length 256;
Best Local Similarity 30.4%; Pred. No. 17;
Matches 7; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 VLGNVIOBNHYDCKXYTKQFN 25
DB 94 IIGSHREKRNHESDEFIAKFN 116

RESULT 15

SC14_CANGA STANDARD; PRT; 302 AA.

AC P53989;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SC14 cytosolic factor (Phosphatidylinositol/phosphatidylcholine
DE transfer protein) (PI/PC TP).
GN SC14.

OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;

RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 350;
RA Dundon W., Ielam K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: REQUIRED FOR TRANSPORT OF SECRETORY PROTEINS FROM THE
CC GOLGI COMPLEX. CATALYZES THE TRANSFER OF PHOSPHATIDYLINOSITOL AND
CC PHOSPHATIDYLCHOLINE BETWEEN MEMBRANES IN VITRO (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE GOLGI COMPLEX AS A
CC PERIPHERAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 CREAL-TRIO domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X97320; CAA65985.1; -.
 DR HSSP; P24280; IADA.
 DR InterPro; IPR001251; CRAL_TRIO.
 DR InterPro; IPR001071; Rebind/tocTrans.
 DR Pfam; PF00650; CRAL_TRIO; 1.
 DR PRINTS; PR03765; CRAL_TRIO.N: 1.
 DR SMART; SM00516; SEC14; 1.
 DR PROSITE; PS50191; CRAL_TRIO; 1.
 KW Transport; Protein transport; Golgi stack.
 FT DOMAIN 97 270
 SQ SEQUENCE 302 AA; 34292 MW; A857E3C70DBA2DDF CRC64;

Query Match 29.4%; Score 45; DB 1; Length 302;
 Best Local Similarity 36.0%; Pred. No. 20;
 Matches 9; Conservative 6; Mismatches 6; Indels 4; Gaps 1;

QY 6 NYIQNPHYDGK---XFTYKQFNK 26
 Db 87 NTIMQDFHYDEKPLVAKYPPQYVHK 111

Search completed: February 10, 2004, 16:55:08
 Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:49:13 / Search time 35 Seconds

(without alignments)
213.815 Million cell updates/sec

Title: US-10-059-447A-1

Sequence: 1 MKVLGVNIQRNFHYDGKFFYTKQFNKPIX 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	82.4	629	16	Q8XNK8 Clostridium
2	56	36.6	2474	12	Q8JUX6 Clostridium
3	54	35.3	687	16	Q8ENJ1 Clostridium
4	53	34.6	510	3	Q8NTU7 Clostridium
5	53	34.6	963	12	Q8Y9U2 Clostridium
6	52.5	34.3	178	8	Q9G9U2 Clostridium
7	52	34.0	512	10	Q8L673 Clostridium
8	52	34.0	513	10	Q9CAT3 Clostridium
9	52	34.0	537	10	Q9JUG1 Clostridium
10	52	34.0	535	10	Q9LYS8 Clostridium
11	52	34.0	750	4	Q8VC40 Clostridium
12	52	34.0	782	4	Q9NVE2 Clostridium
13	52	34.0	785	11	Q8BVN5 Clostridium
14	52	34.0	1392	4	Q9H3T8 Clostridium
15	49.5	32.4	202	15	Q9E4U1 Clostridium
16	49.5	32.4	516	5	O61758 Clostridium

17	49.5	32.4	1050	5	Q9NUN5 Clostridium
18	49.5	32.4	1051	5	Q97123 Clostridium
19	49.5	32.4	1051	8	Q9MTD3 Clostridium
20	49	32.0	42	16	Q8EY83 Clostridium
21	49	32.0	355	16	Q9A8M6 Clostridium
22	49	32.0	374	12	Q9YMP6 Clostridium
23	49	32.0	435	16	Q8DGQ0 Clostridium
24	49	32.0	436	16	Q8Z037 Clostridium
25	48.5	31.7	440	16	Q8U821 Clostridium
26	48	31.4	76	12	Q9YVY4 Clostridium
27	48	31.4	91	16	Q8XXK6 Clostridium
28	48	31.4	126	9	Q9T1F5 Clostridium
29	48	31.4	138	2	Q9JP27 Clostridium
30	48	31.4	179	2	Q8GMJ9 Clostridium
31	48	31.4	185	16	Q33117 Clostridium
32	48	31.4	527	5	P90932 Clostridium
33	48	31.4	815	16	Q9ZKT4 Clostridium
34	48	31.4	1478	5	Q8IKX3 Clostridium
35	48	31.4	1819	12	Q8QEM4 Clostridium
36	48	31.4	2437	12	Q8Q273 Clostridium
37	48	31.4	5922	5	Q8IIN2 Clostridium
38	48	31.4	9271	5	Q8IES7 Clostridium
39	47.5	31.0	788	3	Q8XIN3 Clostridium
40	47	30.7	263	3	Q8NJO0 Clostridium
41	47	30.7	493	8	Q8TID7 Clostridium
42	47	30.7	962	12	Q8Y443 Clostridium
43	47	30.7	1532	5	Q8IC18 Clostridium
44	47	30.7	1591	5	Q9VRP4 Clostridium
45	46.5	30.4	334	2	Q8GKX3 Clostridium

ALIGNMENTS

RESULT 1

ID Q8XNK8 PRELIMINARY, PRT: 629 AA.

AC Q8XNK8; 01-MAR-2002 (TREMREL. 20, Created)

DT 01-MAR-2002 (TREMREL. 20, Last sequence update)

DE 01-OCT-2002 (TREMREL. 22, Last annotation update)

GN CPEB0325 OR AAGA. Hypothetical protein CPEB0325 (Alpha-N-acetylglucosaminidase).

OS Clostridium perfringens.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

NCBI_taxid=1502; Clostridium.

Db 1 MKVLGN1QNRNFHYDGKSFYTTSLNPI 28

RESULT 2

Q8JUX6 PRELIMINARY; PRT; 2474 AA.

AC Q8JUX6; 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Chikungunya virus.
 OS Chikungunya virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 NCBI_TaxID=37124;

RP SEQUENCE FROM N.A.
 RC STRAIN=S27-African prototype;
 RA Khan A.H., Morita K., Parquet M.C., Hasebe F., Mathenge E.G.M.,
 RA Igatahi A.;
 RT "Complete nucleotide sequence of Chikungunya virus and evidence of an
 internal polyadenylation site."
 RL J. Gen. Virol. 0:0-0(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=S27-African prototype;
 RA Khan A.H., Morita K., Parquet M.C., Hasebe F., Mathenge E.G.M.,
 RA Igatahi A.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF369024; AAN05101.1; -

DR InterPro; IPR002589; A1Pp.
 DR InterPro; IPR002620; Peptidase_C9.
 DR InterPro; IPR001788; RNA_dep_RNAp12.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR006066; Viral_helicase1.
 DR Pfam; PF01661; A1Pp; 1.
 DR Pfam; PF01707; Peptidase_C9; 1.
 DR Pfam; PF00978; RNA_dep_RNAp12; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR SMART; SMO0506; A1Pp; 1.
 DR PROSITE; PS50507; RDRP_POSITIVE; 1.
 DR PROSITE; PS50521; RDRP_VIRAL; 1.
 KM Polypeptidn.
 SQ SEQUENCE 2474 AA; 275649 MW; 97D250B9E5A3B0 CRC64;

Query Match 36.6%; Score 56; DB 12; Length 2474;
 Best Local Similarity 47.6%; Pred. No. 40;
 Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 8 IQNPNHYDGKKFYTKQFNKP 28

Db 855 IVSSLHTEGKMTTNEINKPI 875

RESULT 3

Q8EN31 PRELIMINARY; PRT; 687 AA.

AC Q8EN31; 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Catalase (EC 1.11.1.6).
 GN KATB OR O82659.
 OS Oceanobacillus theysensis.
 OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
 NCBI_TaxID=182710;

RP SEQUENCE FROM N.A.
 RC STRAIN=HT831 / DSM 14371 / JCM 11309;
 RX MEDLINE=2220767; PubMed=1235376;
 RA Takami H., Takaki Y., Uchiyama I.;

RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya
 Ridge and its unexpected adaptive capabilities to extreme
 environments."
 RT Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; AP004602; BAC14615.1; -

KM Oxidoreductase; Peroxidase; Complete proteome.
 SQ SEQUENCE 687 AA; 77756 MW; FCCBC8C8267110A0 CRC64;

Query Match 35.3%; Score 54; DB 16; Length 687;
 Best Local Similarity 46.2%; Pred. No. 20;
 Matches 12; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 5 GNYIQNPNHYDGKKFYTKQFN--KPI 28

Db 615 GDHINDQFAVEAGNFTQFNHYKPI 640

RESULT 4

Q8NCU7 PRELIMINARY; PRT; 510 AA.

AC Q8NCU7; 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Gag.
 GN Gag.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10261;
 RA Goodwin T.C.D., Foulter R.T.M.;
 RT "Tc3 and Tcd3: unusual Ty3/gypsy-like retrotransposons from Candida
 albicans and Candida dubliniensis."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF510498; AAM43936.1; -

SQ SEQUENCE 510 AA; 59231 MW; 0BAFB3C72A72A1E CRC64;
 Query Match 34.6%; Score 53; DB 3; Length 510;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 RNFHYDGKKFYTKQFNKP 27

Db 463 RNSNTNPKTFYPRPDKP 480

RESULT 5

Q8V9U2 PRELIMINARY; PRT; 963 AA.

AC Q8V9U2; 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE RNA helicase.
 GN L09CL.
 OS African swine fever virus (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.

NCBI_TaxID=10497;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malawi L11-20/1;
 RA Roberts P.C., Lu Z., Rock D.L.;
 RT "Nucleotide sequence and analysis of 16.25 kilobase pairs of the
 African swine fever virus genome that span the central variable
 region."
 RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; L00966; AAL31320.1; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR002464; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.

DR Pfam, PF00271, helicasase_C.1.
 DR SMART, SM00487, DEXDC, 1.
 DR SMART, SM00490, HELIC, 1.
 DR PROSITE, PS00690, DEAH_ATP_HELICASE, 1.
 KM ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 963 AA; 109831 MW; C1137DAB22E2810F CRC64;

Query Match 34.6%; Score 53; DB 12; Length 963;
 Best Local Similarity 50.0%; Pred. No. 40;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 LGNYIOENFYHYDGKXF 19
 DB 786 LHNIVKTHFHFSGKSP 801

RESULT 6

Q9G902 PRELIMINARY; PRT; 178 AA.
 AC Q9G902;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Ribosomal protein L6.
 GN RPL6.
 OS Ochromonas danica.
 OC Mitochondrion.
 OC Eukaryota; Stramenopiles; Chrysophyceae; Ochromonadales;
 OC Ochromonadaceae; Ochromonas.
 OX NCBI_TaxID=2986;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Burger G., Lang B.F., Gray M.W.;
 RT "Phylogenetic relationships of stramenopile algae, based on complete
 RT mitochondrial genome sequences.";
 RT submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL, AF287134; AAG18403.1;
 DR InterPro, IPR00702; Ribosomal_L6.
 DR Pfam, PF00347; Ribosomal_L6; 1.
 DR PRINTS, PR00059; RIBOSOMALL6.
 DR Prodom, PD002236; Ribosomal_L6; 1.
 KM Mitochondrion.
 SQ SEQUENCE 178 AA; 20956 MW; 8213C3F9D3CD9DF CRC64;

Query Match 34.3%; Score 52.5; DB 8; Length 178;
 Best Local Similarity 38.9%; Pred. No. 8.1;
 Matches 14; Conservative 3; Mismatches 6; Indels 13; Gaps 2;

QY 4 LGNYIOENFYHYDGKXF-----YTKQFNK 26
 DB 142 LGNFIQKIFHLRSYDCYKGGKGFSEFNKIKLKQFK 177

RESULT 7

Q8LG73 PRELIMINARY; PRT; 512 AA.
 AC Q8LG73;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation.";
 RL Genome Biol. 0:0-0(2002).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Broeyer V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana.";
 RT submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY084430; AAM61003.1;
 KM Hypothetical protein.
 SQ SEQUENCE 512 AA; 56168 MW; 34EA672AE3C151F5 CRC64;

Query Match 34.0%; Score 52; DB 10; Length 512;
 Best Local Similarity 53.3%; Pred. No. 29;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 12 FHYDGKRYTKQFNK 26
 DB 388 FNYDGKRYTKQFNK 402

RESULT 8

Q9CAT3 PRELIMINARY; PRT; 513 AA.
 AC Q9CAT3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 56.3 kDa protein.
 GN T18K17.15 OR AT1G73180.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA STRAIN=cv. Columbia;
 RC MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhym T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill U.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miltseher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Vays R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banb J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
 RA Yu S., Bowser L., Cammici P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamliya A., Karlin-Neumann G., Kawai J., Kim C.,
 RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Sarcu M., Seki M., Shinn P.,
 RA Southwick A., Tracy S.E., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene T18K17.15 (GI:12324325).";
 RT submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banb J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamita Y., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyer M.C., Miranda M., Narusaka M., Nguyen M., Nguyen M.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A., Clones.",
 RT "Arabidopsis Open Reading Frame (ORF) Clones.",
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010556; AAC52134.1; -
 DR EMBL; AY056302; AAL07151.1; -
 DR EMBL; AY091307; AAM14246.1; -
 KM Hypothetical protein.
 SQ SEQUENCE 513 AA; 56297 MW; 21B21CE67DE8DAD CRC64;

Query Match 34.0%; Score 52; DB 10; Length 513;
 Best Local Similarity 53.3%; Pred. No. 29;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

12 FHVGGKFFTKQFNK 26
 389 FNDGKRYKKMFER 403

RESULT 9
 Q9SGJ1 PRELIMINARY; PRT; 527 AA.
 AC Q9SGJ1
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative chloroplast nucleoid DNA binding protein.
 GN ATG342980.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosida II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ncv. Columbia;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanden S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Nerman W.C.,
 RA Fraser C.M., Venter J.C.,
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006931; AAD21712.2; -
 DR EMBL; AC006580; AAM15292.1; -
 DR HSSP; P00799; 2AS1.
 DR InterPro; IPR001461; AspproteaseA1.
 DR InterPro; IPR001969; Aspprotease_site.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PRO0792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 527 AA; 58649 MW; B662162BBA6E4EBD CRC64;

Query Match 34.0%; Score 52; DB 10; Length 527;
 Best Local Similarity 58.8%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

3 VLGNYIQRNFR-YDGK 17
 497 IIGNYQQNFIHYDTK 513

RESULT 10
 Q9LYS8 PRELIMINARY; PRT; 535 AA.
 AC Q9LYS8
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein (AT3559080/F17016_130).
 GN F17016_130 OR AT3559080.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosida II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
 RA Rued S., Lemcke K., Mayer K.F.X., Queller F., Salanoubat M.,
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita Y.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Tortum M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamita Y., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Topp M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamita Y., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL163527; CAB86936.1; -
 DR EMBL; AF424562; AAL11556.1; -
 DR EMBL; AY099818; AAM20669.1; -
 DR EMBL; BT000326; AAN15645.1; -
 DR HSSP; P00799; 2AS1.
 DR InterPro; IPR001461; AspproteaseA1.
 DR InterPro; IPR001969; Aspprotease_site.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PRO0792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 535 AA; 59242 MW; 4C20AF38A2B66472 CRC64;

Query Match 34.0%; Score 52; DB 10; Length 535;
 Best Local Similarity 58.8%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

3 VLGNYIQRNFR-YDGK 17
 505 IIGNYQQNFIHYDTK 521

RESULT 11
 Q8VC40 PRELIMINARY; PRT; 750 AA.
 AC Q8VC40
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical 82.9 kDa protein.

```

GN BC021875.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021875; AAH21875.1; -.
DR MGI; MGI:2384286; BC021875.
DR InterPro; IPR00504; RNA_rec_mot.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 750 AA; 82874 MW; 7335ECFDFC2460 CRC64;

Query Match
Best Local Similarity 34.0%; Score 52; DB 11; Length 750;
Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

Qy 1 MKVLGNVYIQNPFHYDGKXF-----YTKQFNKPI 28
Db 488 MKLRGMVYIARQLSFTGVTFKIEEVLSSQSYVMKNKAV 525

RESULT 12
Q9NVB2 PRELIMINARY; PRT; 782 AA.
AC Q9NVB2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FLJ10833 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotaka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto T., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001695; BA91842.1; -.
DR InterPro; IPR00504; RNA_rec_mot.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Hypothetical protein.
FT NON TER 782
SQ SEQUENCE 782 AA; 86258 MW; DOBBD5B8B8E4DF6 CRC64;

Query Match
Best Local Similarity 34.0%; Score 52; DB 4; Length 782;
Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

Qy 1 MKVLGNVYIQNPFHYDGKXF-----YTKQFNKPI 28
Db 524 MKLRGMVYIARQLSFTGVTFKIEEVLSSQSYVMKNKAV 561

RESULT 13
Q8BVN5 PRELIMINARY; PRT; 785 AA.
AC Q8BVN5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CDNA FLJ10833 FTS (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK077133; BAC3636.1; -.
DR NON TER 785
SQ SEQUENCE 785 AA; 86515 MW; C8C7D22310A418C9 CRC64;

Query Match
Best Local Similarity 34.0%; Score 52; DB 11; Length 785;
Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

Qy 1 MKVLGNVYIQNPFHYDGKXF-----YTKQFNKPI 28
Db 524 MKLRGMVYIARQLSFTGVTFKIEEVLSSQSYVMKNKAV 561

RESULT 14
Q9H3T8 PRELIMINARY; PRT; 1392 AA.
AC Q9H3T8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MOP-3.
GN MOP-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Blood;
RA Takayama K., Ukai Y., Fujii Y., Yoshimoto M.;
RT "Molecular and Biological Characterization of a New Nuclear Protein,
RT MOP-3 which is highly expressed in Human Monocytes.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB014772; BAB19784.1; -.
DR InterPro; IPR00504; RNA_rec_mot.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 1392 AA; 154157 MW; F1088016981F09D1 CRC64;

Query Match
Best Local Similarity 34.0%; Score 52; DB 4; Length 1392;
Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

Qy 1 MKVLGNVYIQNPFHYDGKXF-----YTKQFNKPI 28
Db 524 MKLRGMVYIARQLSFTGVTFKIEEVLSSQSYVMKNKAV 561

RESULT 15
Q9E4U1 PRELIMINARY; PRT; 202 AA.
AC Q9E4U1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Leroux C.;
RT "HIV-1 shedding.";

```

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF256394; AAC16055.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 202 AA; 22825 MW; 89BF5A4D87F474BB CRC64;

Query Match 32.4%; Score 49.5; DB 15; Length 202;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY 8 IQRNFHYDGKKFY--TKQFN 25
: ||: ||: ||: ||: ||: ||
Db 104 VMHNFNYGGEFFYCDTTLKFN 124

Search completed: February 10, 2004, 16:55:57
Job time : 39 secs

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 26-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132853.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144683.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145090.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

[illegible]

Sun Feb 15 07:29:53 2004

us-10-059-447a-1.rag

Page 5

PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134944.	PR 28-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 03-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 04-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 05-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 06-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137501.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 09-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 10-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 11-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 12-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 13-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139463.	PR 25-AUG-1999;	99US-0150565.
PR 18-JUN-1999;	99US-0139750.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139753.	PR 27-AUG-1999;	99US-0151065.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151066.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151080.
PR 23-JUN-1999;	99US-0140353.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140354.	PR 31-AUG-1999;	99US-0151438.
PR 24-JUN-1999;	99US-0140695.	PR 01-SEP-1999;	99US-0151930.
PR 28-JUN-1999;	99US-0140823.	PR 07-SEP-1999;	99US-0152363.
PR 29-JUN-1999;	99US-0140921.	PR 10-SEP-1999;	99US-0153070.
PR 30-JUN-1999;	99US-0141287.	PR 13-SEP-1999;	99US-0153758.
PR 01-JUL-1999;	99US-0141842.	PR 15-SEP-1999;	99US-0154018.
PR 02-JUL-1999;	99US-0142154.	PR 16-SEP-1999;	99US-0154039.
PR 06-JUL-1999;	99US-0142390.	PR 20-SEP-1999;	99US-0154779.
PR 06-JUL-1999;	99US-0142803.	PR 22-SEP-1999;	99US-0155139.
PR 09-JUL-1999;	99US-0142920.	PR 23-SEP-1999;	99US-0155486.
PR 12-JUL-1999;	99US-0142977.	PR 24-SEP-1999;	99US-0156559.
PR 13-JUL-1999;	99US-0143542.	PR 28-SEP-1999;	99US-0156458.
PR 14-JUL-1999;	99US-0143624.	PR 29-SEP-1999;	99US-0156596.
PR 15-JUL-1999;	99US-0144005.	PR 04-OCT-1999;	99US-0157117.
PR 16-JUL-1999;	99US-0144085.	PR 05-OCT-1999;	99US-0157753.
PR 16-JUL-1999;	99US-0144086.	PR 06-OCT-1999;	99US-0157865.
PR 19-JUL-1999;	99US-0144325.	PR 07-OCT-1999;	99US-0158029.
PR 19-JUL-1999;	99US-0144331.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144332.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144334.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159329.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159330.
PR 21-JUL-1999;	99US-0144684.	PR 14-OCT-1999;	99US-0159331.
PR 21-JUL-1999;	99US-0144814.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0145086.	PR 18-OCT-1999;	99US-0159638.
PR 22-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0159584.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160741.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160770.
PR 23-JUL-1999;	99US-0145145.	PR 21-OCT-1999;	99US-0160814.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160915.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.

25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 34.0%; Score 52; DB 21; Length 512;
 Best Local Similarity 53.3%; Pred. No. 26;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 12 FHYDGKRYTKQFNK 26
 |||||:|:|:|:
 Db 368 FNYDGKRYTKQFNR 402

RESULT 4
 AAY29513
 ID AAY29513 standard; Protein; 705 AA.

AC AAY29513;
 XX
 DT 13-OCT-1999 (first entry)

DE Human lung tumour protein SAL-33 predicted amino acid sequence.

KM Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
 KW immunotherapy; detection; inhibition.

XX Homo sapiens.

PN WO938973-A2.

XX PD 05-AUG-1999.

XX PF 26-JAN-1999; 99WO-US01642.

XX PR 22-DEC-1998; 98US-0219245.

XX PR 28-JAN-1998; 98US-0015022.

XX PR 28-JAN-1998; 98US-0015022.

XX PR 18-MAR-1998; 98US-0040828.

XX PR 18-MAR-1998; 98US-0040831.

XX PR 23-JUL-1998; 98US-0122191.

XX PR 23-JUL-1998; 98US-0122192.

XX PA (CORI-) CORIXA CORP.

XX PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;

XX DR WPI; 1999-479187/40.

XX DR N-PSDB; AAZ07193.

XX PT Lung tumour specific polynucleotides for inhibiting the development

XX PS Claim 3; Page 150-151; 171pp; English.

XX CC The present invention describes lung tumour specific polynucleotides

CC and tumour antigens; AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325

CC represent specifically claimed polynucleotides; and AAY29486 to AAY29571

CC represent amino acid sequences from the present invention. The lung

CC tumour specific polynucleotides and polypeptides can be used in

CC pharmaceutical compositions and vaccines to inhibit the development of

CC lung cancer. They can also be used to detect lung cancer in a patient.

CC Probes and antibodies derived from the lung tumour sequences are useful

CC in detection of lung cancer.

XX Sequence 705 AA;

Query Match 34.0%; Score 52; DB 20; Length 705;
 Best Local Similarity 31.6%; Pred. No. 36;
 Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

QY 1 MKVLGNYIQRNTHYDQKF-----YTKQFNKPI 28
 |||||:|:|:|:
 Db 451 MKLRGMYIARQLSFTGVTFKIEVLLSOSVYVMNKAV 488

RESULT 5
 AAB44462
 ID AAB44462 standard; Protein; 705 AA.

AC AAB44462;

XX DT 05-FEB-2001 (first entry)

DE Human lung tumour-specific antigen encoded by cDNA #99.

XX KM Lung tumour protein; lung cancer; cytostatic; vaccine.

XX OS Homo sapiens.

XX PN WO200060077-A2.

XX PD 12-OCT-2000.

XX PF 30-MAR-2000; 2000WO-US08560.

XX PR 02-APR-1999; 99US-0285323.

XX PR 09-AUG-1999; 99US-0370838.

XX PR 30-DEC-1999; 99US-0476235.

XX PR 03-MAR-2000; 2000US-0518809.

XX PA (CORI-) CORIXA CORP.

XX PI Reed SG, Lodes MJ, Mohamath R, Secrist H;

XX DR WPI; 2000-638466/61.

XX DR N-PSDB; AAC79146.

XX PT Novel lung tumor polypeptides and polynucleotides, useful for

XX detecting, monitoring or treating cancer, especially lung cancer -

XX Example 4; Page 165-167; 243pp; English.

XX CC The present sequence is given in a specification relating to compounds

XX for therapy and diagnosis of lung cancer. Polypeptides comprising at

XX least an immunogenic part of a lung tumour protein are disclosed.

XX CC The polypeptides are useful for inhibiting the development of cancer,

XX CC especially lung cancer. Samples of T cells expressing the polypeptides

XX CC may be used to inhibit the development of cancer. The polypeptides are

XX CC also useful for detecting and monitoring the progression of cancer,

XX CC especially lung cancer.

XX SO Sequence 705 AA;

Query Match 34.0%; Score 52; DB 21; Length 705;
 Best Local Similarity 31.6%; Pred. No. 36;
 Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

QY 1 MKVLGNYIQRNTHYDQKF-----YTKQFNKPI 28
 |||||:|:|:|:
 Db 451 MKLRGMYIARQLSFTGVTFKIEVLLSOSVYVMNKAV 488

RESULT 6
 AAE13803
 ID AAE13803 standard; Protein; 705 AA.

AC AAE13803;
 XX
 DT 26-FEB-2002 (first entry)

XX DE Human lung tumour-specific protein SAL-33.
 XX KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
 XX KW antisense-therapy; vaccine; immune response; lung cancer; SAL-33.
 OS Homo sapiens.
 XX PM WO200172295-A2.
 XX PD 04-OCT-2001.
 XX PF 28-MAR-2001; 2001WO-US09991.
 XX PR 29-MAR-2000; 2000US-0538037.
 XX PR 05-JUN-2000; 2000US-0588937.
 XX PR 18-AUG-2000; 2000US-0640878.
 XX PR 22-SEP-2000; 2000US-234517P.
 XX PR 01-NOV-2000; 2000US-0704512.
 XX PR 14-DEC-2000; 2000US-0738973.
 XX PA (CORI-) CORIXA CORP.
 XX PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
 XX PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
 XX DR WPI: 2001-639201/73.
 XX DR N-PsDB; AAD23221.
 PT New human lung-specific polynucleotides and polypeptides for the
 PT diagnosis and treatment of disease e.g. lung cancer -
 PS Example 4; Page 218-220; 378pp; English.
 CC The invention relates to isolated lung tumour-specific proteins and
 CC their corresponding cDNA molecules. Lung tumour-specific proteins and
 CC their antigen-presenting cells are useful for stimulating and/or
 CC expanding T cells specific for a tumour protein, and for inhibiting
 CC the development of cancer. The invention also relates to a composition
 CC useful for stimulating an immune response, and for treating cancer. The
 CC lung tumour specific oligonucleotide is useful in gene therapy and for
 CC diagnosis, detection and treatment of lung cancer. The present sequence
 CC is human lung tumour-specific protein.
 XX SQ Sequence 705 AA;
 Query Match 34.0%; Score 52; DB 22; Length 705;
 Best Local Similarity 31.6%; Pred. No. 36;
 Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;
 QY 1 MKVTAGNYIQNFHYDGKFF-----YTKQFNKPI 28
 Db 451 MKLRGMYIARQLSFTGVTFKIEEVLISQSYKMKAKAV 488
 RESULT 7
 ID AAB93229 standard; Protein; 782 AA.
 XX AAB93229;
 XX AC AAB93229;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:12223.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX PM EPI074617-A2.
 XX PM 07-FEB-2001.
 XX PD

PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI: 2001-318749/34.
 XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
 XX PT full-length cDNAs defined in the specification, and for the detection
 XX PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 12223; 2537bp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 XX CC full-length cDNAs defined in the specification. Where a primer set
 XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 XX CC to the complementary strand of a polynucleotide which comprises one of
 XX CC the 5602 nucleotide sequences defined in the specification, where the
 XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 XX CC of an oligonucleotide comprising a sequence complementary to the
 XX CC complementary strand of a polynucleotide which comprises a 5'-end
 XX CC sequence and an oligonucleotide comprising a sequence complementary to a
 XX CC polynucleotide which comprises a 3'-end sequence, where the
 XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
 XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
 XX CC the specification. The primer sets can be used in antisense therapy and
 XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 XX CC particularly full-length cDNAs. The primers are also useful for the
 XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
 XX CC the full-length cDNAs. The primers allow obtaining of the full-length
 XX CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 XX CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to
 XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 XX CC represent oligonucleotides, all of which are used in the exemplification
 XX CC of the present invention.
 XX SQ Sequence 782 AA;
 Query Match 34.0%; Score 52; DB 22; Length 782;
 Best Local Similarity 31.6%; Pred. No. 41;
 Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;
 QY 1 MKVLGNYIQNFHYDGKFF-----YTKQFNKPI 28
 Db 524 MKLRGMYIARQLSFTGVTFKIEEVLISQSYKMKAKAV 561
 RESULT 8
 ID AAB68943 standard; Protein; 910 AA.
 XX AAB68943;
 XX AC AAB68943;
 XX DT 20-JAN-2003 (first entry)
 XX DE Human polypeptide SEQ ID NO 990.
 XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 XX KW cell-proliferative disorder; neurodegenerative diseases; bacterial;
 XX KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 XX KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 XX KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
 XX KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 XX KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;

KW antiarthritic.
 XX Homo sapiens.
 OS WO200270539-A2.
 FN 12-SEP-2002.
 PD 05-MAR-2002; 2002WO-US05095.
 PF 05-MAR-2001; 2001US-0799451.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI, 2002-759812/82.
 DR N-PSDB; ABZ11160.
 XX New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX Claim 9; SEQ ID NO 990; 1012pp + Sequence Listing; English.
 PS The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, infection,
 CC ulcer, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 910 AA;
 SQ Query Match 34.0%; Score 52; DB 23; Length 910;
 Best Local Similarity 31.6%; Pred. No. 48;
 Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;
 QY 1 MKVLGNTYQNFHNDGKGF-----YTKQFNKRI 28
 DB 525 MKLKGNTIAROLSTGVTFKIEVVLSSQSYKMNKAV 562
 RESULT 9
 ABP6229
 ID ABP6229 standard; Protein; 765 AA.
 XX
 AC ABP6229;
 XX
 DT 19-MAY-2003 (first entry)
 XX
 DE Human nucleic acid associated protein 12 SEQ ID NO:12.
 XX
 KM Human: nucleic acid associated protein; NMAP; cardiac; cytosolic;
 KM neuroprotective; gene therapy; cardiovascular disorder; cancer;
 KM neurological disorder.
 XX
 OS Homo sapiens.
 PR

PN WO2003016549-A2.
 XX 27-FEB-2003.
 PD 14-AUG-2002; 2002WO-US25829.
 PF 17-AUG-2001; 2001US-313111P.
 PR 24-AUG-2001; 2001US-314682P.
 PR 24-AUG-2001; 2001US-314756P.
 PR 27-AUG-2001; 2001US-315105P.
 PR 31-AUG-2001; 2001US-316751P.
 PR 31-AUG-2001; 2001US-316856P.
 PR 05-OCT-2001; 2001US-328185P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Barroso I, Baughn MR, Becha SD, Blake UT, Borowsky ML, Burford N;
 PI Dugan BM, Elliott VS, Emerling BM, Foreythe ID, Gietzen KJ,
 PI Gorvad AE, Griffin JA, Hatalia AJA, Honchell CD, Ison CH, Khan FA;
 PI Lal PG, Lee EA, Lee SY, Li JX, Lu DM, Lu Y,
 PI Lehr-Mason PM, Nguyen DB, Ramkumar J, Sprague WM, Tang YT;
 PI Thangavelu K, Thornton M, Tran UK, Wallia NK, Warren BA, Xu Y;
 PI Yao MG, Yue H, Yue H, Zebardjian Y;
 DR WPI, 2003-248159/25.
 DR N-PSDB; ABZ79884.
 XX
 PT New human nucleic acid-associated proteins polypeptide, useful for
 PT preparing a composition for diagnosing or treating e.g., cardiovascular
 PT or neurological disorders -
 XX Claim 1; Page 216-217; 290pp, English.
 PS ABZ79873 to ABZ79905 encode the human nucleic acid associated proteins
 CC designated NMAP-1 to NMAP-33 given in ABP6218 to ABP6250. The NMAP
 CC sequences have cardiac, cytosolic and neuroprotective activities, and
 CC can be used in gene therapy. The NMAP sequences can be used for
 CC preparing a composition for diagnosing or treating a disease or
 CC condition associated with decreased expression or overexpression of
 CC functional NMAP e.g. cardiovascular or neurological disorders or cancer.
 XX
 SQ Sequence 765 AA;
 QY Query Match 32.7%; Score 50; DB 24; Length 765;
 Best Local Similarity 44.4%; Pred. No. 79;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 9 QNFHYDGKXPYTKQFNK 26
 DB 209 QPNYYTKLKNFYKQYHK 226
 RESULT 10
 ABB55253
 ID ABB55253 standard; Protein; 307 AA.
 XX
 AC ABB55253;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein cfa.
 XX
 KM Lactococcus lactis IL1403.
 OS
 XX
 PN FR2807446-A1.
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX WPI; 2002-043418/06.
 XX New nucleotide sequence useful in the identification or Lactococcus
 XX lactis and related species -
 XX Claim 6; SEQ ID No 1955; 2504pp; French.
 XX The present invention is related to a Lactococcus lactis nucleotide
 XX sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The
 XX nucleic acid sequence is useful in the detection and/or amplification of
 XX nucleic acid sequence, particularly to identify Lactococcus lactis or
 XX related species. The proteins of the invention are useful for the
 XX biosynthesis or biodegradation of a composition of interest. The
 XX invention helps research in lactic bacteria, particularly useful in the
 XX production of yogurt and cheese.
 XX Note: The sequence data for this patent is based on equivalent patent
 XX WO200177334 (published 18-OCT-2001) which is available in electronic
 XX format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
 XX Sequence 307 AA;
 SQ
 Query Match 32.4%; Score 49.5; DB 23; Length 307;
 Best Local Similarity 41.4%; Pred. No. 35;
 Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 2;
 QY 1 MKVLGNYIQR-----NPHVD-GKGFYTK 22
 Db 101 IKLGNHSQKSRSDISHYDINDFYNK 129
 RESULT 11
 ABP73883
 ID ABP73883 standard; Protein: 567 AA.
 AC ABP73883;
 XX 30-JAN-2003 (first entry)
 XX Candida albicans essential protein SEQ ID NO 7720.
 DE Fungus: Yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;
 KW proliferation; Candida albicans; fungicide; antifungal.
 OS Candida albicans.
 XX WO200253728-A2.
 PN 11-JUL-2002.
 PD 26-DEC-2001; 2001WO-US49486.
 PF 29-DEC-2000; 2000US-259128P.
 PR 20-FEB-2001; 2001US-0792024.
 PR 22-AUG-2001; 2001US-314050P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 XX WPI; 2002-566694/60.
 DR N-PSDB; AB232433.
 XX Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele
 PT of a gene and placing other allele of the gene under conditional
 XX expression -

PS Claim 44; SEQ ID NO 7720; 167pp + Sequence Listing; English.
 XX The invention relates to constructing (M1) a strain of diploid fungal
 XX cells in which both alleles of a gene are modified, comprising modifying
 XX one allele by insertion or replacement by a cassette having an
 XX expressible selectable marker and modifying other allele by
 XX recombination, of a promoter replacement fragment with a heterologous
 XX promoter, so that expression of the second allele is regulated by the
 XX promoter. (M1) is useful for constructing a strain of diploid fungal
 XX cells in which both alleles of a gene are modified. The diploid fungal
 XX cells having both alleles modified are useful for identifying a gene that
 XX is essential to the survival or growth of a fungus, a gene that
 XX contributes to the virulence and/or pathogenicity of a fungus, a gene
 XX that contributes to the resistance of a diploid fungus to an antifungal
 XX agent, an antifungal agent that inhibits the growth of a diploid fungus
 XX and for identifying a therapeutic agent for treatment of a mammalian
 XX disease. (M1) is useful for identifying a compound which modulates the
 XX activity of a gene product, preferably enzymatic activity, carbon
 XX compound catabolism, signal transduction, DNA replication and cell division
 XX translational, signal transduction, transporter, transcriptional,
 XX compound catabolism, signal transduction, DNA replication and cell division
 XX activity. The method is useful for identifying a compound having the
 XX ability to inhibit growth or proliferation of C. albicans cells and for
 XX treating infection by C. albicans. The present sequence is that of an
 XX essential Candida albicans protein used in the method of the invention.
 XX Note: The sequence data for this patent is not represented in the printed
 XX specification but is based on sequence information supplied to Derwent by
 XX the European Patent Office.
 SQ Sequence 567 AA;
 Query Match 32.0%; Score 49; DB 23; Length 567;
 Best Local Similarity 69.2%; Pred. No. 80;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 14 YDGKKFYTKQFVK 26
 Db 278 YPGKLFYTRQFIK 290
 RESULT 12
 AA000958
 ID AA000958 standard; Protein: 110 AA.
 XX AA000958;
 AC AA000958;
 XX 06-NOV-2001 (first entry)
 XX Human polypeptide SEQ ID NO 14850.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorders; arthritis; inflammation.
 OS Homo sapiens.
 XX WO200164835-A2.
 PN 07-SEP-2001.
 PD 26-FEB-2001; 2001WO-US04927.
 PF 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR N-PSDB; AA180889.
 XX Isolated nucleic acids and polypeptides, useful for preventing


```

RESULT 15
AAV11066
ID AAV11066 standard; Protein; 815 AA.
XX
AC AAV11066;
XX
DT 08-JUN-1999 (first entry)
XX
DE H. pylori ORF 13ae10610_156411_c3_33 outer membrane protein.
XX
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
XX secreted protein; cytoplasmic protein; cellular protein.
OS Helicobacter pylori.
XX
PN WO9824475-A1.
XX
PD 11-JUN-1998.
XX
PE 05-DEC-1997; 97WO-US22104.
XX
PR 14-JUL-1997; 97US-0891928.
XX 05-DEC-1996; 96US-0759625.
PR 25-MAR-1997; 97US-0823745.
XX
PA (ASTRA ) ASTRA AB.
XX
PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;
XX
DR WPI; 1998-333051/29.
DR N-PSDB; AAX30595.
XX
PT New isolated Helicobacter pylori nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of infection by
PT H. pylori and other Helicobacter species
XX
PS Claims 37, 41; Page 228-230; 339pp; English.
XX
CC Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 97 ORFs are shown. The proteins are variously cell envelope proteins,
CC cytoplasmic proteins, secreted proteins or other cellular proteins.
CC Vaccines containing the nucleic acids or proteins are claimed, as are
CC probes containing at least 8 nucleotides from the nucleic acid
CC sequences. The vaccines are useful for treating or reducing the risk of
CC H. pylori infections, and the probes can be used diagnostically for
CC detecting the presence of Helicobacter in a sample. The products are
CC also of use in screening for compounds having the ability to interfere
CC with the H. pylori life cycle or to inhibit H. pylori infection.
XX
SQ Sequence 815 AA;
XX
Query Match 31.4%; Score 48; DB 19; Length 815;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 9; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 3 VLGNVIOBNFHYDGKFFTKQ 23
Db 199 IMGAY--RNHHFDALVYTHQ 217

```

Search completed: February 10, 2004, 16:54:49
 Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:51:12 ; Search time 22 Seconds

(without alignments)
55.773 Million cell updates/sec

Title: US-10-059-447A-1

Sequence: 1 MKVLGNLYIQRNHYDGKXFKYTKQFNKPI 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUTS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	98.0	29	4	US-09-185-476B-1
2	52	34.0	705	4	US-09-370-838-186
3	47	30.7	80	4	US-09-732-210-353
4	46.5	30.4	1312	2	US-08-592-126-148
5	46.5	30.4	1312	2	US-08-687-080-51
6	46.5	30.4	1312	4	US-09-158-595-148
7	46.5	30.4	1341	4	US-09-252-991A-26785
8	46	30.1	541	4	US-08-867-941-14
9	46	30.1	541	3	US-09-074-658-14
10	46	30.1	1297	4	US-09-107-532A-4552
11	46	30.1	2439	3	US-09-074-658-11
12	45	29.4	1319	3	US-08-290-731C-2
13	45	29.4	1333	3	US-08-356-952-7
14	45	29.4	1336	2	US-08-230-731C-6
15	45	29.1	1752	2	US-09-865-621A-2
16	44	28.8	80	4	US-09-252-991A-29905
17	44	28.8	98	1	US-08-375-346A-2
18	44	28.8	98	3	US-08-467-123B-2
19	44	28.8	98	3	US-08-943-336A-2
20	44	28.8	98	4	US-09-635-899-2
21	44	28.8	98	5	PCT-US95-01780-2
22	44	28.8	181	4	US-09-134-001C-5665
23	44	28.8	201	4	US-09-134-001C-5559
24	44	28.8	2431	1	US-07-920-281C-2
25	44	28.8	2431	3	US-08-466-277-2
26	43	28.1	80	4	US-09-732-210-344
27	43	28.1	119	4	US-08-858-207A-503

28	43	28.1	156	2	US-09-070-060-7	Sequence 7, Appli
29	43	28.1	156	3	US-09-051-969A-3	Sequence 3, Appli
30	43	28.1	156	3	US-09-051-969A-4	Sequence 4, Appli
31	43	28.1	156	3	US-09-357-746-7	Sequence 7, Appli
32	43	28.1	384	4	US-08-530-862B-4	Sequence 4, Appli
33	43	28.1	384	4	US-08-597-313D-4	Sequence 4, Appli
34	43	28.1	566	4	US-09-134-001C-3431	Sequence 3431, Ap
35	43	28.1	1356	1	US-08-810-116-8	Sequence 8, Appli
36	43	28.1	1356	2	US-07-930-548A-8	Sequence 8, Appli
37	43	28.1	1356	2	US-09-098-707A-2	Sequence 2, Appli
38	43	28.1	1356	4	US-09-483-539-2	Sequence 4, Appli
39	43	28.1	1367	1	US-07-813-593-4	Sequence 4, Appli
40	43	28.1	1367	1	US-07-977-451-6	Sequence 6, Appli
41	43	28.1	1367	1	US-07-946-507-4	Sequence 4, Appli
42	43	28.1	1367	1	US-08-252-517-6	Sequence 6, Appli
43	43	28.1	1367	1	US-07-906-397A-6	Sequence 6, Appli
44	43	28.1	1367	1	US-08-601-891-6	Sequence 6, Appli
45	43	28.1	1367	2	US-08-443-861-2	Sequence 2, Appli

ALIGNMENTS

```

RESULT 1
US-09-185-476B-1
; Sequence 1, Application US/09185476B
; Patent No. 6399749
; GENERAL INFORMATION:
; APPLICANT: Smith, Dan
; TITLE OF INVENTION: ALPHA-N-ACETYLGALACTOSAMINIDASE FROM CLOSTRIDIUM PERFRINGENS
; FILE REFERENCE: 0994.00093
; CURRENT APPLICATION NUMBER: US/09/185,476B
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/064,683
; PRIOR FILING DATE: 1997-11-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Clostridium perfringens
; FEATURE:
; NAME/KEY: Xaa
; LOCATION: (1)..(29)
; OTHER INFORMATION: Xaa is any amino acid
US-09-185-476B-1
Query Match          98.0%; Score 150; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MKVLGNLYIQRNHYDGKXFKYTKQFNKPI 28
DB      1 MKVLGNLYIQRNHYDGKXFKYTKQFNKPI 28
RESULT 2
US-09-370-838-186
; Sequence 186, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE OF INVENTION: LONG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289

```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 186
LENGTH: 705
TYPE: PRT
ORGANISM: Homo sapiens
US-09-370-838-186

Query Match 34.0%; Score 52; DB 4; Length 705;
Best Local Similarity 31.6%; Pred. No. 5.8;
Matches 12; Conservative 4; Mismatches 12; Indels 10; Gaps 1;

QY 1 MKVLGNVYQNFHYDGKXFFYTKQFNKPI 28
DB 451 MKLNGVYIARQLSFTGVTFKLEIVLLSQSYVKMYNKAV 488

RESULT 3
US-09-732-210-353
Sequence 353; Application US/097322210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Miltanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Mu, Yornie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1/53
SEQ ID NO: 80
LENGTH: 80
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-732-210-353

Query Match 30.7%; Score 47; DB 4; Length 80;
Best Local Similarity 45.8%; Pred. No. 3.2;
Matches 11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 4 LGNVYQNFHYDGKXFFYTKQFNKPI 27
DB 56 MGSHPDR--HYCGKCLITCFNKP 77

RESULT 4
US-08-592-126-148
Sequence 148; Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Rad50, pro-translation of SEQ ID NO: 54
US-08-592-126-148

Query Match 30.4%; Score 46.5; DB 2; Length 1312;
Best Local Similarity 46.9%; Pred. No. 82;
Matches 15; Conservative 1; Mismatches 7; Indels 9; Gaps 2;

QY 1 MKVLGNVYQNFHYDGKXFFYTKQ---FNKPI 28
DB 958 MKDIENYIQ-----GKDKYKQKETEINKVI 984

RESULT 5
US-08-687-080-51
Sequence 51; Application US/08687080
Patent No. 5965427
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human Rad50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: TRANS. OF RAD50 CDNA (SEQ. 54), NT.
INDIVIDUAL ISOLATE: 389 TO 4324
US-08-667-080-51

Query Match 30.4%; Score 46.5; DB 2; Length 1312;
Best Local Similarity 46.9%; Pred. No. 82;
Matches 15; Conservative 1; Mismatches 7; Indels 9; Gaps 2;

QY 1 MKVIGNYIQRFPHDQKXFTYKQ---FNKPI 28
Db 958 MKDIENYIQ-----DGKDYKKXKXETELNKVI 984

RESULT 6
US-09-168-595-148
Sequence 148; Application US/09168595
Patent No. 6555866

GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/168,595
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,126
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Rad50, pro-translation of SEQ ID NO:54
US-09-168-595-148

Query Match 30.4%; Score 46.5; DB 4; Length 1312;
Best Local Similarity 46.9%; Pred. No. 82;
Matches 15; Conservative 1; Mismatches 7; Indels 9; Gaps 2;

QY 1 MKVIGNYIQRFPHDQKXFTYKQ---FNKPI 28
Db 958 MKDIENYIQ-----DGKDYKKXKXETELNKVI 984

RESULT 7
US-09-252-991A-26785
Sequence 26785; Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26785
LENGTH: 1341
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26785

Query Match 30.4%; Score 46.5; DB 4; Length 1341;
Best Local Similarity 29.6%; Pred. No. 84;
Matches 8; Conservative 9; Mismatches 7; Indels 3; Gaps 1;

QY 5 GNYIQRFPHDQKXFTYKQ---FNKPI 28
Db 363 GVIVSENFYDSEIVDTKIVAFTRPL 389

RESULT 8
US-08-867-941-14
Sequence 14; Application US/08867941
Patent No. 5977337

GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M
APPLICANT: Du, Run-Pan
APPLICANT: Wang, Qitjun
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H
TITLE OF INVENTION: LACTOPERRICIN RECEPTOR GENES OF MORAXELLA
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,941
FILING DATE: 03-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-681 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 541 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-867-941-14

Query Match 30.1%; Score 46; DB 2; Length 541;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 GNYIORNFHYDGKXFY 20
|||:| |||:|
Db 393 GNTYERYQGDGKRKH 408

RESULT 9
US-09-074-658-14
; Sequence 14, Application US/09074658
; Patent No. 6184371
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Run-Pan Du
; APPLICANT: Quljun Wang
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,658
; FILING DATE: 08-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-074-658-14

Query Match 30.1%; Score 46; DB 3; Length 541;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 GNYIORNFHYDGKXFY 20
|||:| |||:|
Db 393 GNTYERYQGDGKRKH 408

RESULT 10
US-09-107-532A-4552
; Sequence 4552, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariandello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4552:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1297
; SEQUENCE DESCRIPTION: SEQ ID NO: 4552:
US-09-107-532A-4552

Query Match 30.1%; Score 46; DB 4; Length 1297;
Best Local Similarity 60.0%; Pred. No. 97;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 KVLGNVYIORNFHYDG 16
|||:| |||:|
Db 101 KVLGNVYIORNFHYDG 115

RESULT 11
US-09-074-658-11
; Sequence 11, Application US/09074658
; Patent No. 6184371
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Run-Pan Du
; APPLICANT: Quljun Wang
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,658
; FILING DATE: 08-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-795
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1153
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2439 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-074-658-11

Query Match 30.1%; Score 46; DB 3; Length 2439;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 GNYIQRNFHYDQKXFTQ 20
DB 2291 GNYTERYQGYDQKRYH 2306

RESULT 12
US-08-290-731C-2
Sequence 2, Application US/08290731C
Patent No. 5843646
GENERAL INFORMATION:
APPLICANT: BOWTELL, David Douglas Lawrence
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE
TITLE OF INVENTION: SON OF SEVENLESS (MSOS) GENE,
TITLE OF INVENTION: AND MSOS POLYPEPTIDES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,731C
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00068
FILING DATE: 17-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PL0921/92
FILING DATE: 17-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: 0-36066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-731C-2

Query Match 29.4%; Score 45; DB 2; Length 1319;
Best Local Similarity 39.1%; Pred. No. 1.4e+02;

Matches 9; Conservative 6; Mismatches 4; Indels 4; Gaps 1;
QY 1 MKVLGNYIQRNFHYDQKXFTQ 23
DB 142 LKLVGNVYVNRIRHYE----ITKQ 160

RESULT 13
US-09-356-952-2
Sequence 2, Application US/09356952
Patent No. 6117663
GENERAL INFORMATION:
APPLICANT: Bortack-Sjodin, Ann
APPLICANT: Margalit, S. M.
APPLICANT: Bor-Soggi, Dafna
APPLICANT: Cole, Philip
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 600-1-228N
CURRENT APPLICATION NUMBER: US/09/356,952
CURRENT FILING DATE: 1999-07-19
EARLIER APPLICATION NUMBER: 60/093,631
EARLIER FILING DATE: 1998-07-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1333
TYPE: PRT
ORGANISM: Homo sapiens
US-09-356-952-2

Query Match 29.4%; Score 45; DB 3; Length 1333;
Best Local Similarity 39.1%; Pred. No. 1.4e+02;
Matches 9; Conservative 6; Mismatches 4; Indels 4; Gaps 1;

QY 1 MKVLGNYIQRNFHYDQKXFTQ 23
DB 142 LKLVGNVYVNRIRHYE----ITKQ 160

RESULT 14
US-08-290-731C-6
Sequence 6, Application US/08290731C
Patent No. 5843646
GENERAL INFORMATION:
APPLICANT: BOWTELL, David Douglas Lawrence
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE
TITLE OF INVENTION: SON OF SEVENLESS (MSOS) GENE,
TITLE OF INVENTION: AND MSOS POLYPEPTIDES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,731C
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00068
FILING DATE: 17-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PL0921/92
FILING DATE: 17-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: KIT, Gordon
 REGISTRATION NUMBER: 30,764
 REFERENCE/DOCKET NUMBER: Q-36066
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1336 amino acids
 TYPE: amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-290-731C-6

Query Match

29.4%; Score 45; DB 2; Length 1336;

Best Local Similarity 39.1%; Pred. No. 1.4e+02;

Matches 9; Conservative 6; Mismatches 4; Indels 4; Gaps 1;

QY

1 MKVLGNYIQRFHYDGKFFTKQ 23

Db

159 LKLVENVVNRIRHYE---ITKQ 177

RESULT 15

US-09-865-621A-2

Sequence 2, Application US/09865621A

Patent No. 6553348

GENERAL INFORMATION:

APPLICANT: JORGENSEN, FLEMING

APPLICANT: HANSEN, OLE CAI

APPLICANT: STORGARD, PETER

TITLE OF INVENTION: A NEW ENZYME ISOLATED FROM A BIFIDOBACTERIUM

FILE REFERENCE: 02405.0200

CURRENT APPLICATION NUMBER: US/09/865,621A

CURRENT FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: 60/207,154

PRIOR FILING DATE: 2000-05-26

NUMBER OF SEQ ID NOS: 41

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 1752

TYPE: PRT

ORGANISM: Bifidobacterium bifidum

US-09-865-621A-2

Query Match

29.1%; Score 44.5; DB 4; Length 1752;

Best Local Similarity 47.4%; Pred. No. 2.3e+02;

Matches 9; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY

10 RNFHYDGKXFTKQPNKPI 28

Db

921 RSFY-SRNYVYKTKGNKPI 938

Search completed: February 10, 2004, 16:56:48
 Job time : 42 secs


```

; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 446
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-310-154-446

Query Match          32.7%; Score 50; DB 12; Length 510;
Best Local Similarity 52.9%; Pred. No. 37;
Matches 9; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY      1 MKVLGNVIOQNFHYDGGKFFYTKQFNKP 27
Db      478 MSIIQNFQGNFHYVD 494

RESULT 6
US-10-369-493-18892
; Sequence 18892, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18892
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(430)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-369-493-18892

Query Match          32.0%; Score 49; DB 12; Length 430;
Best Local Similarity 33.3%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY      1 MKVLGNVIOQNFHYDGGKFFYTKQFNKP 27
Db      29 MAYVANDVETTYKQSDPFFYLTGFNEB 55

RESULT 7
US-10-032-585-7720
; Sequence 7720, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

```

```

; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7720
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Candida albicans
; US-10-032-585-7720

Query Match          32.0%; Score 49; DB 12; Length 567;
Best Local Similarity 69.2%; Pred. No. 58;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      14 YDGGKFFYTKQFNKP 26
Db      278 YPGKLFYTRQPIK 290

RESULT 8
US-10-369-493-5186
; Sequence 5186, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5186
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-10-369-493-5186

Query Match          31.4%; Score 48; DB 12; Length 527;
Best Local Similarity 33.3%; Pred. No. 75;
Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY      2 KVLGNVIOQNFHYDGGKFFYTKQFNKP 28
Db      381 KVDRHINRNPLNKETLFEKSDNKP 407

RESULT 9
US-10-393-602-148
; Sequence 148, Application US/10393602
; Publication No. US20030170714A1
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Delinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/393,602
;   FILING DATE: 19-Mar-2003
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/08/592,126
;     FILING DATE: 26-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sholtz, Charles K.
;     REGISTRATION NUMBER: 38,615
;     REFERENCE/DOCKET NUMBER: 4600-0111
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 324-0880
;     TELEFAX: (415) 324-0960
;
; INFORMATION FOR SEQ ID NO: 148:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1312 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHEICAL: NO
;   ORIGINAL SOURCE:
;     INDIVIDUAL ISOLATE: Rad50, pro-translation of SEQ ID NO:54
;   SEQUENCE DESCRIPTION: SEQ ID NO: 148:
US-10-393-602-148

Query Match          30.4%; Score 46.5; DB 12; Length 1312;
Best Local Similarity 46.9%; Pred. No. 3.3e+02;
Matches 15; Conservative 1; Mismatches 7; Indels 9; Gaps 2;

QY 1 MKVLGNVIGRNFFHYDGKXFYTKQ---FNKPI 28
   ||:|||||:|||||:|||||:|||||
Db 958 MKDIENYIQ-----DGKDDYKKQKXETLNKVI 984

RESULT 10
US-10-369-493-19944
; Sequence 19944, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;   APPLICANT: Cao, Yongwei
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   APPLICANT: Goldman, Barry S.
;   APPLICANT: Chen, Xianfeng
;   TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;   FILE REFERENCE: 38-10(52052)B
;   CURRENT APPLICATION NUMBER: US/10/369,493
;   CURRENT FILING DATE: 2003-02-28
;   PRIOR APPLICATION NUMBER: US 60/360,039
;   PRIOR FILING DATE: 2002-02-21
;   NUMBER OF SEQ ID NOS: 47374
;   SEQ ID NO 19944
;   LENGTH: 584
;   TYPE: PRT
;   ORGANISM: No. US20030233675A1loc punctiforme
US-10-369-493-19944

Query Match          30.1%; Score 46; DB 12; Length 584;
Best Local Similarity 42.9%; Pred. No. 1.6e+02;
Matches 9; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 5 GNVIGNRHHYDGKXFYTKQPN 25
   :|||:|||||:|||||:|||||
Db 415 GRVVSANFSYDGKQIITGSGN 435

RESULT 11
US-09-764-891-4533
; Sequence 4533, Application US/09764891
; Publication No. US20030077808A1
```

```

; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC006
;   CURRENT APPLICATION NUMBER: US/09/764,891
;   CURRENT FILING DATE: 2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 10231
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 4533
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-891-4533

Query Match          29.7%; Score 45.5; DB 11; Length 147;
Best Local Similarity 44.4%; Pred. No. 45;
Matches 12; Conservative 6; Mismatches 4; Indels 5; Gaps 3;

QY 6 NYIGRNHY-DG--KXRYTKQFN--KP 27
   :||:|||||:|||||:|||||
Db 61 HYDEKNWYCDGYDRRFYETIINGLKP 87

RESULT 12
US-10-264-049-3481
; Sequence 3481, Application US/10264049
; Publication No. US2004005579A1
; GENERAL INFORMATION:
;   APPLICANT: Bize et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PA133PI
;   CURRENT APPLICATION NUMBER: US/10/264,049
;   CURRENT FILING DATE: 2002-10-04
;   PRIOR APPLICATION NUMBER: PCT/US01/18569
;   PRIOR FILING DATE: 2001-06-07
;   PRIOR APPLICATION NUMBER: US 60/209,467
;   PRIOR FILING DATE: 2000-06-07
;   NUMBER OF SEQ ID NOS: 4360
;   SOFTWARE: Patentin Ver. 3.1
;   SEQ ID NO 3481
;   LENGTH: 163
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: MISC FEATURE
;     LOCATION: (136)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;   FEATURE:
;     NAME/KEY: MISC FEATURE
;     LOCATION: (138)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;   FEATURE:
;     NAME/KEY: MISC FEATURE
;     LOCATION: (141)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;   FEATURE:
;     NAME/KEY: MISC FEATURE
;     LOCATION: (148)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3481

Query Match          29.4%; Score 45; DB 12; Length 163;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 13 HYDGKXFYTKQFNKP 27
   |||:|||||:|||||:|||||
Db 146 HYXGKWLCTYWFNKP 160

RESULT 13
US-10-264-237-1704
```



```
; Sequence 1704, Application US/10264237
; Publication No. US2004009491A1
; GENERAL INFORMATION:
; APPLICANT: Birex et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1704
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (81)-
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (241)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (291)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; US-10-264-237-1704

Query Match          29.4%; Score 45; DB 12; Length 308;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 GNYIQRN---FHYDG 16
DB 109 GNYIELNENDPHYNG 124

RESULT 14
; US-09-867-550-1188
; Sequence 1188, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Rued,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1188
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-867-550-1188
```

```
Query Match          29.4%; Score 45; DB 9; Length 318;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 GNYIQRN---FHYDG 16
DB 183 GNYIELNENDPHYNG 198

RESULT 15
; US-09-764-864-842
; Sequence 842, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 842
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-864-842

Query Match          29.4%; Score 45; DB 10; Length 422;
Best Local Similarity 56.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 5 GNYIQRN---FHYDG 16
DB 159 GNYIELNENDPHYNG 174

Search completed: February 10, 2004, 17:01:31
Job time : 35 secs
```